

Inspection with Robotic Microscopic Imaging

Liam Pedersen, Matthew Deans, Clay Kunz
QSS Group Inc, NASA Ames Research Center, Moffett Field, CA, 94035

Randy Sargent
Carnegie Mellon West, NASA Ames Research Center, Moffett Field, CA 94035

Alan Chen
Stanford University, Palo Alto, CA, 94305

and

Greg Mungas
Jet Propulsion Laboratory, Pasadena, CA, 91109

Abstract

Future Mars rover missions will require more advanced onboard autonomy for increased scientific productivity and reduced mission operations cost. One such form of autonomy can be achieved by targeting precise science measurements to be made in a single command uplink cycle. In this paper we present an overview of our solution to the subproblems of navigating a rover into place for microscopic imaging, mapping an instrument target point selected by an operator using far away science camera images to close up hazard camera images, verifying the safety of placing a contact instrument on a sample or finding nearby safe points, and analyzing the data that comes back from the rover. The system developed includes portions used in the Multiple Target Single Cycle Instrument Placement demonstration at NASA Ames in October 2004, and portions of the MI Toolkit delivered to the Athena Microscopic Imager Instrument Team for the MER mission still operating on Mars today. Some of the component technologies are also under consideration for MSL mission infusion.

I. Introduction

MICROSCOPIC imagers are valuable tools for rover based science and engineering tasks, from studying small scale morphology of soils, rocks, and potential biota to the inspection of equipment. Rover operations and data analysis with a microscopic imager present unique challenges that are not generally addressed in the robotics literature. Typically, a microscopic imager must be brought very close to a feature of interest in order to get a high magnification image. Doing so from a distance requires bulky optics unsuited to a rover payload. Near field imaging requires driving a rover up to a target feature while precisely keeping track of its position relative to the vehicle. Bringing an arm mounted camera lens close to an uneven surface like a rock presents a risk, particularly in unstructured and partially unknown environments. Images from a microscope typically have a limited depth of field, so that many images at different focal lengths are required in order to get focused imagery of the entire field of view. This can require repositioning of the camera between images, leading to rotations, translations and scale changes. Image processing techniques for creating a focused composite image must account for this.

We have developed methods for our robotic vehicle, K9, to autonomously navigate to multiple rock features scattered within a 10m diameter area, and deploy a microscopic imager against them, all in a



Figure 1: The K9 rover is comparable in size to the MER rovers, with steerable cameras on a mast and fixed cameras overlooking the 5DOF arm workspace.

single command cycle. This represents a tenfold improvement, as measured by the number of features that can be investigated close up, over MER class vehicles which require 3 command cycles, each lasting a single sol, to approach and investigate a *single* feature. The system uses a vision-based target tracker that recovers the 6-DOF transformations between the rover and the tracked targets as the rover moves. The tracker is comprised of a feature based approach that tracks a set of interest points in 3-D using stereo, and a shape based approach that registers dense 3-D meshes. Autonomous analysis of close up stereo models of rocks enables safe instrument placement on the target feature.

In addition, we developed and delivered the *MER MI Toolkit*, to support the microscopic imager (MI) instrument on the MER mission. The MI Toolkit is a set of routines to register MI images to create focused composite images, MI mosaics and 3D models. These tools are in use by the MI instrument team in MER science operations and have also been ported to the CHAMP microscopic imager used on the K9 rover at NASA Ames.

The paper is organized as follows. Section II presents some related work. Section III discusses the onboard vision methods used for tracking targets during navigation and mapping the selected point to rover imagery in the final rover position. Section IV presents a method for analyzing potential placement locations for instrument safety. Section V describes ground based MI and CHAMP data processing tools, and finally Section VI offers some discussion of the work presented.

II. Related work

There is strong interest in science autonomy and instrument placement capabilities for planetary rover applications. Recently, Wettergreen et al. have focused on autonomy for kilometer long traverses and investigation scenarios where search is less structured[33]. Alternatively, robotics groups at both JPL and NASA Ames have been focusing on shorter range and higher precision science activity, developing autonomy for machine vision, navigation, positioning, and instrument placement for precisely defined targets meters away from a rover. Tools such as Viz[34] provide interfaces for designating science targets, and tools exist for determining rover navigation goal positions in order to visit those science targets[1]. Ground based planning and onboard execution generate and carry out the operations plans that best satisfy a rich set of prioritized goals and resource constraints[26]. JPL's visual odometry[24] is designed to significantly increase navigation precision. The technique has been validated in terrestrial analog field tests and demonstrated in flight on MER. Ames has developed a similar method to precisely track the location of science targets while navigating[6] focusing instead on the precise location of multiple science targets relative to the rover. Several other visual tracking techniques have been developed or evaluated specifically for precision navigation to science targets with surface rovers[2][17][21][7]. Precision manipulation for planetary rover arms has also been addressed[22], including the current state-of-the-art instrument arm positioning for the MER rovers[3]. Much of this development is done within, or using, the CLARAty software framework[32]. The image analysis tools described in Section V were developed for the MER Microscopic Imager (MI) instrument, whose development was led by USGS[12][13], and later ported to the CHAMP instrument, developed by the LASP laboratory at the University of Colorado in Boulder[19].

III. Vision Based Tracking and Approach

Localization errors from rover odometry and deduced reckoning are too large to guide a rover to a small scale target over large distances with the required accuracy. Therefore, the rover must explicitly track target locations as it navigates about the worksite and avoids obstacles. Because features are selected for scientific relevance, they are not necessarily those features which best facilitate visual tracking. The rover might move completely around targets, causing self occlusion of features of interest. Lighting may change due to shadows cast by the rover or changing sun angles over the course of a Sol.



Figure 2: CHAMP (Camera Hand-lens And Microscope Probe) microscopic camera, a MIDP developed instrument with 6um/pixel resolution and a movable CCD detector array for obtaining a “z-stack” of images at different focal lengths.

Many visual feature based trackers operate by matching a chosen template to an area of interest in successive images. The search is often done using an exhaustive correlation or convolution, which can be expensive when precise predictions are not available or large camera motions must be accommodated. In addition, some trackers offer the user the flexibility to specify a template, but the specified template may not be amenable to tracking due to low visual texture or changing appearance during motion. In addition, if the tracker only keeps track of one nominal target point, it is brittle in the event of a mismatch, and vulnerable to occlusions, changing viewpoints or other real world effects.

The appearance based tracking algorithm used in our system uses large numbers of image features matched across stereo pairs. Feature detection and matching is done automatically using the SIFT algorithm[20], which consists of an interest point detector to find salient points in images, a descriptor to summarize the appearance of detected points, and a matching algorithm to find similar descriptors. SIFT typically finds hundreds to thousands of points in each of the rover cameras, and matches around 25% to 50% of the features with only a small number of outliers, typically about 1% to 3%. Our 3D SIFT based tracker uses these matched features to recover the motion of the tracked target. SIFT provides matched pairs of image points $z_i^{(t)} = (l_i^{(t)}, r_i^{(t)})$ and $z_j^{(t+1)} = (l_j^{(t+1)}, r_j^{(t+1)})$ from left and right images at two discrete time steps. Calibrated stereo is used to recover the 3D locations $x_i^{(t)}$ and $x_j^{(t+1)}$ of the points. We then estimate the 6-DOF transformation T that aligns one view to the next by minimizing

$$\sum_{ij \in c} (x_j^{(t+1)} - T^{(t+1)} x_i^{(t)})^2$$

using Horn's rotation fitting algorithm[14] and RANSAC[8]. Horn's method finds the optimal transformation in closed form, but due to the second order cost function, outliers in matching can cause arbitrarily large errors in the recovered transformation. To identify and eliminate outliers we use RANSAC to find the transformation that is consistent with the largest number of inliers. Inliers are defined as those matches $(x_i^{(t)}, x_j^{(t+1)})$ such that

$$|x_j^{(t+1)} - T^{(t+1)} x_i^{(t)}| < \tau$$

where τ is a threshold. Currently we use $\tau = 3$ cm and repeat the RANSAC loop 100 times using 3 putative matches in each trial, which takes negligible computation time. RANSAC returns the transformation with the largest consensus, and the list of matches in the consensus set. To further improve the estimate we use the consensus set to re-estimate the transform with all of the inliers. Once the rigid transformation $T^{(t+1)}$ is computed, the tracked feature location is simply updated by applying the transformation to the target location.

Updates occur after each meter of rover motion. The tracking algorithm typically tracks targets with an accuracy of 1 mm per meter of motion, or around 0.1% distance traveled.

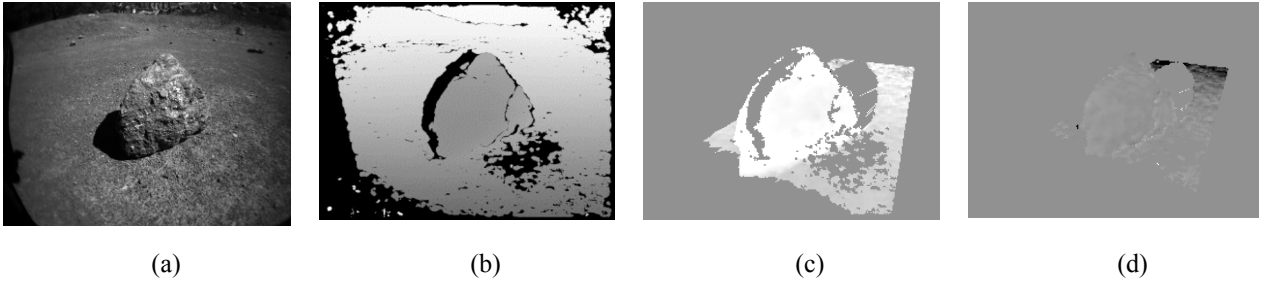


Figure 3: Registration result. (a) Hazcam view of rock with rover in place. (b) Depth map from hazcam stereo. (c) Depth errors for initial guess at alignment. (d) Final depth errors after alignment.

Even with this precision, visual tracking alone is insufficient for placing the instrument at the specified target with high precision. Our method corrects for errors in tracking by registering the initial view, in which the rover operator selected the target point, to the final view, available to the rover at its terminal position once the target has been approached.

The initial and final stereo image pairs are used to construct 3D models of the target. The rigid transformation that aligns the two models can be used to determine the coordinate transformation between views. This transformation can be used to map the initial designated point to the final view for precise placement. Our mesh registration approach projects these two models into a virtual range sensor view and minimizes the difference between the rendered depths at each point. The 3D models v and v' are represented by triangulated

meshes. For each triangle on the mesh, the vertices v_i , v_j , and v_k are projected onto the image plane to find the bounding region. Then for every pixel in the bounding triangle, the location of the intersection of the camera ray c and the facet of the mesh is a point s_i , given by

$$s_i = a_i v_i + a_j v_j + a_k v_k$$

with $a_i + a_j + a_k = 1$. The depth to the intersection point is the z coordinate in the camera frame,

$$z_i = n_c^T s_i$$

which is the projection of the point s_i onto the normalized camera ray n_c . The vector of all depths z_i is denoted z , and the range image under transformation parameterized by p is denoted $z(p)$. The cost function to be minimized is a function of the difference in range images

$$J(p) = \sum_i f(z_i(p) - z_i')$$

The surface model v' does not move during registration, so z' is a constant. The depth to the point v_i changes with transformation, so the rendering operation is done for each trial solution during optimization. The rendering takes $O(n)$ operations, where n is the number of pixels in the virtual range sensor. In order to accommodate outliers in stereo model building, we use a robust norm[27] for $f()$.

To improve the optimization of $J(p)$ under local minima, we first perform a coarse correlation search in order to initialize the search close to the global optimum. Correlation over 6 dimensions is prohibitive, but we make a few approximations to limit the search to 2 dimensions. The transformation is estimated as part of the feature based tracker above, and rover orientation is measured directly to within a few degrees by onboard orientation sensors. Using the observed orientation reduces the search space to 3 dimensions. Since we are minimizing a difference in depth images, we perform a correlation in the 2 dimensions parallel to the image plane. If there is an average difference in depth, it can be computed directly and subtracted out.

Once the correlation search finds an approximate solution, we optimize over all 6 rigid transformation parameters using Nelder-Mead[27], which is a general local nonlinear optimization method. Nelder-Mead only requires a cost function, not any derivative information, so the cost function is used directly. In order to avoid problems with early termination, we restart the Nelder-Mead optimization twice after it converges. Figure 5 shows an example result of the depth error after convergence.

The transformation estimated by the 3D registration step describes how the original view and final view align. The same transformation is applied to the selected target point to find the same point in the final view for instrument placement.

IV. Instrument Safety Check

Robotic manipulation of a contact sensor such as the CHAMP in unstructured environments requires conservative checks for instrument safety prior to actual placement against the target. At peak magnification the CHAMP imager has a working distance of only a few millimeters, which is near the limit of positioning information available via proprioception. CHAMP has three contact switches, and for peak magnification the instrument arm positions the camera near the surface and then drives the camera along its normal until the contact switches close. Rough surfaces, large protrusions or holes, or edges of objects can damage the instrument.

These instrument safety checks need to be done when the robot is close up to the target feature. Besides the fact that operators may not be able to determine if a target point is safe from 10m distance, there is no certainty that the rover will have tracked the target point with sufficient precision to avoid placing the instrument on an adjacent unsafe zone. Without automated safety check and arm motion planning, an additional command cycle is required for each target.

Our instrument safety check uses several heuristics to verify that the selected imaging target is safe, or find a safe alternative nearby. The method consists of thresholding several statistics computed from a 3D model of the target in which all points are given priority levels. The highest priority is the chosen target point, and the priority of alternate points decreases with distance from the target. Points are then checked in priority order. The heuristics use a set of parameters to reflect different constraints for different instruments.

The first check is for surface roughness. A plane is fit to the points within some radius of the point being evaluated, where the radius corresponds to the circle circumscribing the tool interface. The plane fit yields three

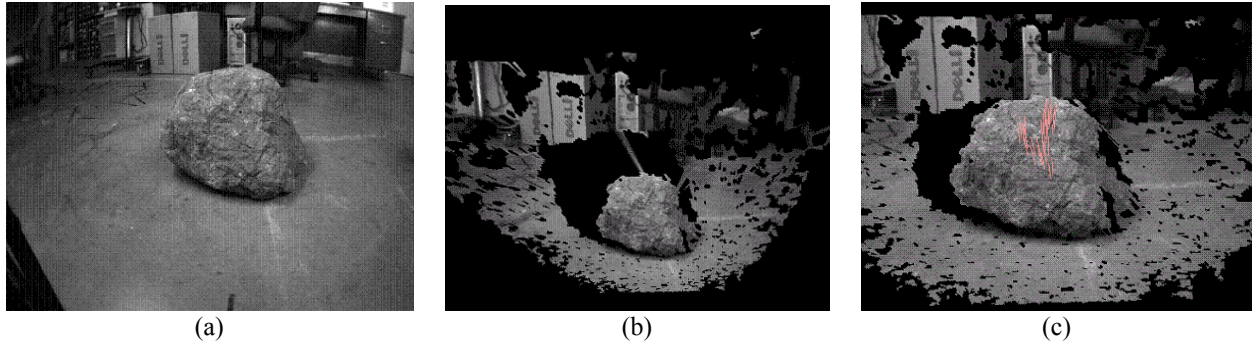


Figure 4: This 3D model is the basis for 3D shape registration for precise matching of the selected point, as well as analysis for instrument safety. (a) Hazcam view of a rock. (b) 3D model from hazcam stereo. (c) some points which have passed the instrument safety check, with the tool normal shown.

statistics. The first is the surface normal. If the angle between the viewing direction and surface normal is too large, the point is rejected as too oblique. The second is the residual to the fit, or surface roughness. If the average deviation from a plane is too large, the region is rejected as too rough. The third statistic is the maximum deviation

from the planar fit. If the maximum deviation is too large the region is rejected. The tool bounding radius, maximum obliquity, maximum roughness, and maximum deviation are parameters which can be set for each tool.

The algorithm also looks at the percentage of valid stereo correspondences near the point of interest. If the tool bounding radius contains an occlusion boundary or a textureless region, then stereo may fail often and the point is rejected because not enough is known about the shape. The minimum percentage of valid stereo correspondence is also a parameter that can be set by the designer. These heuristics are shown in **Figure 5**.

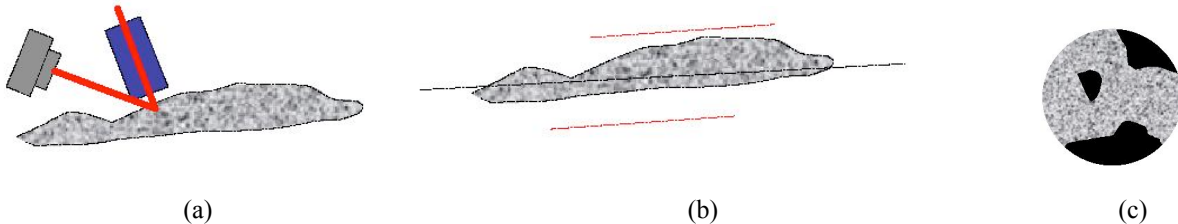


Figure 5: Instrument safety checks. (a) surface normal. (b) surface roughness and maximum deviation. (c) valid stereo coverage and maximum hole size.

V. Data Analysis

Near field imagers with reduced depth of field and arm mounted monocular imagers require data analysis tools that differ from those used for other rover imagers such as mast mounted navigation camera pairs. The MI Toolkit was developed to automatically perform image registration and *focal section merging*, the combining several images of a surface into a single maximally in-focus image, for the Athena Microscopic Imager instrument team.

A. Image Registration

Focal section merging requires that we first find corresponding pixels in a stack of MI images. For the best focal section merging, we want a dense, subpixel estimate for correspondences, which can be expensive to compute when image motion is large and there are no *a priori* epipolar constraints. In order to facilitate the search for these correspondences, we first register the images up to a homography. If the scene is relatively planar, this homography accounts for most of the image plane motion of scene points, and the job of the pixel by pixel correspondence search is made much simpler.

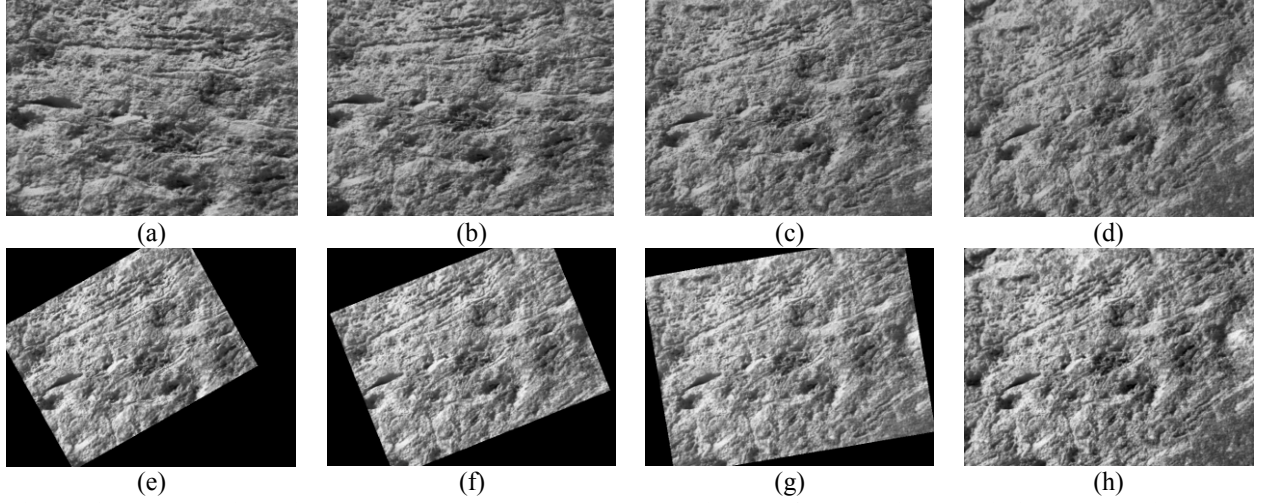


Figure 6: Image registration example. Top row shows source images from a microscopic imager moving away from a target surface and rotating. Bottom row (e-h) shows images aligned to the reference view.

The image registration uses a nonlinear optimization method to recover the 8 parameters describing the homography. The homography H describes the relationship between pixel coordinates in images I_1 and I_2 as

$$I_1(x) = I_2(x') = I_2(Hx) \quad (1)$$

where x and x' are projective coordinates in 2D, and the equality is up to a scale factor. The matrix H describing the homography has the form

$$H = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \quad (2)$$

We choose the lower right entry to be unity because the homography is only defined up to a scale factor. We then search for the other 8 parameters. In addition, the two images may have slightly different exposures, which we account for using a linear function of the pixel values, and each pixel will contain some random noise, so that equation (1) becomes

$$I_1(x) = \alpha I_2(Hx) + \beta + \varepsilon_x \quad (3)$$

For convenience, the 8 parameters of the homography and the linear coefficients α and β can be collected into the parameter vector $\theta = (a, b, c, d, e, f, \alpha, \beta)$. We now define a least squares cost function over θ ,

$$J(\theta) = \sum_x [I_1(x) - \alpha I_2(Hx) + \beta]^2 \quad (4)$$

and minimize $J(\theta)$ using Levenberg-Marquardt.

Working with full resolution images presents two problems. First, the cost function in (4) may have lots of local minima which can trap the minimization. Second, the derivatives and cost function can take a long time to compute. For both of these reasons, we use a coarse-to-fine approach based on image pyramids.

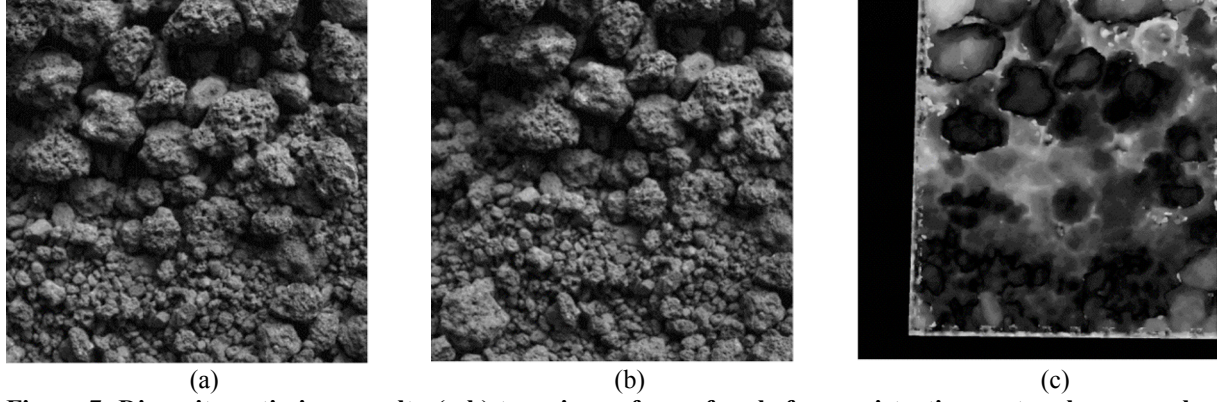


Figure 7: Disparity optimizer result. (a-b) two views of a surface before registration up to a homography. (c) magnitude of disparity corrections at each pixel, showed as greyscale. Black corresponds to 0 pixels, white to approximately 23 pixels in this example.

The image pyramid is constructed for both I_1 and I_2 up to level 3, or subsampling to 2^{-3} the original image dimensions. These subsampled images are aligned only up to rotation and scale, i.e.

$$H_3 = \begin{bmatrix} \cos(\phi) & -\sin(\phi) & u \\ \sin(\phi) & \cos(\phi) & v \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

The level 3 registration starts with H_3 initialized to the identity matrix. Once the level 3 registration converges, the level 2 images are registered, again up to rotation and scale. In order to initialize the search at level 2, we start with the rotation angle recovered at level 3 and multiply the translation parameters by 2 to account for the difference in scale at the next pyramid level. The same procedure followed at level 1, and the final registration is a search for the full homography at level 0 (full resolution), initialized by H_1 .

Figure 7 shows an example image sequence using an engineering model of the MER MI taken at Cornell University, as well as the result of the registration.

B. Dense Correspondence

A stack of images registered using homography is a useful product. Among other things it allows scientists to browse a registered stack, scrolling from near to far focus, without the distraction of large image motions. However, homography does not correct for parallax due to camera motion relative to a non-planar surface.

Although parallax is typically associated with lateral camera motion, it is also present for motion along the camera pointing direction, increasing towards the edges of the image. Parallax motions of several pixels are commonly observed between successive MI images, while even a fractional pixel is enough to cause artifacts when performing the focal section merge.

We initially used a simple window correlation search to find correspondences between images. For a given pixel in the first image, a pixel is chosen in the second image minimizing the sum square of differences of corresponding pixels in a window around each pixel.

Correcting each image in a stack in this way, we created a set of images we hoped would exhibit no relative motion when viewed sequentially, only changes in focus. Disappointingly, this wasn't the case. Small motions (less than a pixel) were apparent when animating between images, and this miscorrelation caused features to grow or shrink in the resulting focal section merge.

We first modified the window correlation search to calculate a subpixel match location by modeling the neighborhood of the optimal match as a quadratic basin and finding its minimum. This improved the results, but unfortunately small motions persisted in areas of the image around large depth relief.

The continued problem turned out to be the parallax itself; different portions of the support window around a pixel were moving by different amounts. Depending on the texture of the scene, the area very close to the pixel might dominate the match solution, or an area towards the edge of the window, which moved differently. Minimizing the support window size reduced this effect at the cost of spurious incorrect solutions and noise in the quadratic subpixel solution.

To solve our problem, we developed a new technique for maintaining support window size while removing the problem of parallax. We used an iterative approach to solving the correspondences. For the first iteration, a complete correlation search is performed between the two images, which will be called A and B . This correlation provides our initial estimate for parallax motion. For each successive iteration, we use the parallax estimation from the previous iteration to warp image A into image A' , and then perform a correlation between A' and B . This correlation provides a correction to our current estimate of parallax motion.

As the iterations continue, A' approaches B in appearance, and as it does, the error in correlation minimum due to parallax reduces. The resulting solution produces images without perceptible subpixel motion, allowing us to generate focal section merges without perceptible feature size changes. **Figure 7** shows the magnitude of the parallax correction at each pixel after homography.

C. Focal Section Merging

The motivation for the development of the MI toolkit is the problem of focal section merging, which is necessary because of the limited depth-of-field of the MI's optics. While focus is a good cue for relative depth for humans as they flip back and forth between images in a vertical sequence, it is also useful to be able to examine a single, globally in-focus image of a non-planar scene.

After a stack is captured by the MI and a dense correspondence is found, images are transformed using the dense correspondence to create a new stack of images, where each pixel x,y on one image corresponds directly to the same pixel x,y on all other images. Next, these transformed images are combined into a single, in-focus image, called the focal section merge.

A simple approach to combining the images would be to compute each pixel in the focal section merge by selecting the pixel with sharpest focus from that location in the stack of images transformed to remove motion. Unfortunately, the transformed images have undergone interpolation from the original images, destroying some of the focus information. Our approach is, for each pixel in the transformed image, to use the inverse of the dense correspondence map, and use the corresponding location in the original image to compute sharpness..

Sharpness of focus $s(x,y)$, is calculated from local variation in image intensity $i_{x,y}$:

$$s(x,y) = \sum_{\substack{x'=x-3\sigma^2 \\ y'=y-3\sigma^2}}^{\substack{x'=x+3\sigma^2 \\ y'=y+3\sigma^2}} e^{-\frac{(x-x')^2+(y-y')^2}{2\sigma^2}} \left((i_{x',y'} - i_{x'-1,y'})^2 + (i_{x',y'} - i_{x',y'-1})^2 \right) \quad (5)$$

$$\sigma^2 = 3$$

After these sharpness values are computed, the focal section merge is constructed using the pixel from each location with the highest sharpness value. Figure 8 shows two images, followed by the focal section merge.

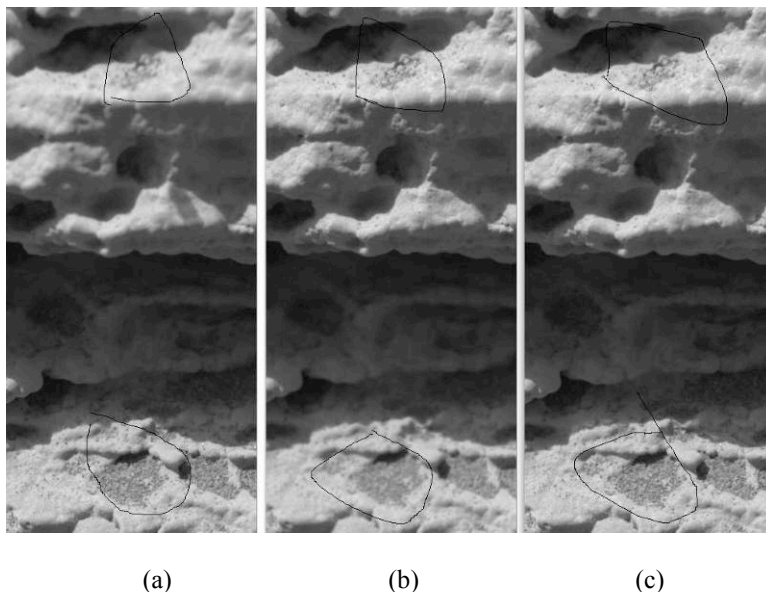


Figure 8: A focal section merge uses focused sections of multiple images (a-b) to assemble a new, fully focused view of a sample (c).

VI. Conclusion

The overarching goal of NASA's Mars Exploration Program is to answer the question "Did life ever exist on Mars?" If it did, or if it still does, expectations are that it would likely be microscopic. Microscopy is therefore an essential tool for detecting and characterizing extinct or extant life or viable habitats in any detail. It is also fundamental to geology. The Microscopic Imager (MI) on the Mars Exploration Rovers (MER) has provided dramatic illustration of the critical importance of microscopy in the exploration of Mars's surface, particularly when mobility is available, so that a variety of geologic sites may be explored. For example, the MER MI played an instrumental role in the discovery and analysis of finely layered rippled bedforms and blueberry shaped concretions that helped confirm the existence of shallow aqueous and salt rich environments in the Martian past.

Because life in extreme environments is both rare and heterogeneously distributed, finding it requires investigating many locations, diversely distributed at both macroscopic and microscopic scales. Thorough analysis of candidate features is necessary to unambiguously detect life and draw meaningful conclusions. In regions where small fractions (i.e. 0.1%) of potential microhabitats actually harbor life, the amount of activity required to carry out a meaningful tele-robotic search goes beyond the current demonstrated state-of-the-art even for terrestrial analog robotic capabilities. When the risks, latencies, and data bandwidth constraints of planetary surface operations are added, the productivity of robotic science is minimized.

The 2009 Mars Science Laboratory (MSL) offers the next opportunity to conduct extensive close up analyses of Martian rocks from a mobile platform after the MER missions. The MSL mission scenario assumes MER level instrument placement capabilities and calls for intensive, long duration (5 sols or more) analyses of Martian rocks to search for potential microhabitats.

We have successfully demonstrated a complete integrated rover system capable of safely acquiring contact measurements from at least 4 distinct rocks, scattered over a 10m diameter area, in a single command cycle lasting less than a day. Whilst work remains to validate the system and port to flight relevant computational hardware, this represents a tenfold increase in MER class rover capability, as measured by number of features investigated, and would increase MSL science productivity by at least 30%.

More importantly, the capability to rapidly get close up microscopic image mosaics at various resolutions of features scattered over an area, using a microscopic imager and other short duration measurements, enables a more

aggressive strategy of sample triage to identify promising targets much more effectively than can be done with remotely acquired measurements alone, and greatly improving the odds that the necessary, exhaustive and time consuming analysis of select samples will yield results. The Ames Single Cycle Instrument Placement (SCIP) system and MI Toolkit are essential components towards realizing this capability to effectively explore extreme planetary environments with robotic vehicles.

Acknowledgements

The authors would like to acknowledge the support of many of the Intelligent Robotics Group, the CLARAty development team, and the Athena MI Instrument Team for supporting software, rover hardware, and operations. We would also like to thank Michael Sims and Ken Herkenhoff for their guidance on image analysis for the MI instrument. This work was supported by the Intelligent Systems, Collaborative Decision Systems, and Mars Technology programs.

References

- [1] P.Backes, "Automated Rover Positioning and Instrument Placement", in *IEEE Aerospace 2005*, Big Sky, Montana, March 5-12, 2005.
- [2] Bajracharya, M., A. Diaz-Calderon, M. Robinson, M. Powell, "Target Tracking, Approach and Camera Handoff for Automated Instrument Placement," in *IEEE Aerospace 2005*, Big Sky, Montana, March 5-12, 2005.
- [3] Baumgartner, E., R.G. Bonitz, J.P. Melko, L.R. Shiraishi, P.C. Leger, "The Mars Exploration Rover Instrument Positioning System," in *IEEE Aerospace 2005*, Big Sky, Montana, March 5-12, 2005.
- [4] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239-256, 1992.
- [5] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. In *IEEE International Conference on Robotics and Automation*, volume 3, pages 2724-2729, 1991.
- [6] M. Deans, C. Kunz, R. Sargent, E. Park, and L. Pedersen, "Combined feature based and shape based visual tracker for robot navigation," in *IEEE Aerospace 2005*, Big Sky, Montana, USA, March 2005.
- [7] M. Deans, C. Kunz, R. Sargent, L. Pedersen, "Terrain Model Registration for Single Cycle Instrument Placement," in *International Symposium on Artificial Intelligence and Robotics in Space (iSAIRAS)*, Nara, Japan, 2003.
- [8] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381-395, June 1981.
- [9] A. Fitzgibbon. Robust registration of 2d and 3d point sets. In *British Machine Vision Conference*, pages 411-420, 2001.
- [10] D. Gennery. Calibration and Orientation of Cameras in Computer Vision, chapter Least-Squares Camera Calibration Including Lens Distortion and Automatic Editing of Calibration Points, pages 123-136. Springer Verlag (A. Gruen and T. Huang, ed.), 2001.
- [11] C. Harris and M. Stephens. A combined corner and edge detector. In *Fourth Alvey Vision Conference*, pages 147-151, 1988.
- [12] K.E. Herkenhoff et al., Athena Microscopic Imager investigation, *J. Geophys. Res.*, 108(E12), 8065, doi: 10.1029/2003JE002076, 2003.

- [13] K.E. Herkenhoff, et al., Textures of the Soils and Rocks at Gusev Crater from Spirit's Microscopic Imager. *Science*, Vol. 305:824--826, 2004.
- [14] B. K. P. Horn. *Robot Vision*. MIT Press, 1986.
- [15] T. Huntsberger, H. Aghazarian, Y. Cheng, E.T. Baumgartner, E. Tunstel, C. Leger, A. Trebi-Ollennu, P.S. Schenker, "Rover Autonomy for Long Range Navigation and Science Data Acquisition on Planetary Surfaces," in *IEEE Int. Conf. on Robotics and Automation*, Washington, D.C., May 2002.
- [16] Y. Ke and R. Sukthankar. Pca-sift: A more distinctive representation for local image descriptors. In *Computer Vision and Pattern Recognition*, 2004.
- [17] W. Kim, R. Steinke, R. Steele, "2-D Target Tracking Technology Validation Report," JPL Technical Report D-28523, Apr. 2004.
- [18] W.S.Kim, A.I.Ansar, R.D.Steele, "Rover Mast Calibration, Exact Camera Pointing, and Camera Handoff for Visual Target Tracking," in *IEEE International Conference on Robotics and Automation (ICRA)*, May 2005.
- [19] G.M. Lawrence, J.E. Boynton, and et al. Champ: Camera handlens microscope. In *The 2nd MIDP Conference, Mars Instrument Development Program*, 2000.
- [20] D. G. Lowe. Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, 60(2):91-110, 2004.
- [21] I. Nesnas, M. Bajracharya, R. Madison, E. Bandari, C. Kunz, M. Deans, M. Bualat, .Visual Target Tracking for Rover-based Planetary Exploration., *IEEE Aerospace Conference*, Big Sky, Montana, 2004.
- [22] I. Nesnas, M. Maimone, H. Das, .Autonomous Vision- Based Manipulation from a Rover Platform, *IEEE Symp. on Computational Intelligence in Robotics and Automation*, 1999.
- [23] L. Nguyen, M. Bualat, L. Edwards, L. Flueckiger, C. Neveu, K. Schwehr, M. Wagner, E. Zbinden, "Virtual reality interfaces for visualization and control of remote vehicles," *Autonomous Robots* 11(1), 2001.
- [24] C. F. Olson, L. H. Matthies, M. Schoppers, and M. W. Maimone. Robust stereo ego-motion for long distance navigation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 453-458, 2000.
- [25] Eric Park, Linda Kobayashi, and Susan Y. Lee. Extensible hardware architecture for mobile robots. In *IEEE International Conference on Robotics and Automation*, 2005, under submission.
- [26] L. Pedersen, M. Bualat, D.E. Smith, R. Washington, "Integrated Demonstration of Instrument Placement, Robust Execution and Contingent Planning," *International Symposium on Artificial Intelligence and Robotics in Space (i-SAIRAS)*, Nara, Japan, 2003.
- [27] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1988.
- [28] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151-172, 2000.
- [29] J. Shi and C. Tomasi. Good features to track. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 593-600, 1994.

- [30] C. Stoker, E. Zbinden, T. Blackmon, B. Kanefsky, J. Hagen, C. Neveu, D. Rasmussen, K. Schwehr and M. Sims, "Analyzing Pathfinder Data Using Virtual Reality and Superresolved Imaging", *Journal of Geophysical Research - Planets* Vol 104(E4) pages 8889-8906, April 25, 1999.
- [31] C. Urmson, R. Simmons, and I. Nesnas. A generic framework for robotic navigation. In *Proceedings of the 2003 IEEE Aerospace Conference*, 2003.
- [32] R. Volpe, et al. The CLARAty architecture for robotic autonomy. In *Proceedings of the 2001 IEEE Aerospace Conference*, 2001.
- [33] D. Wettergreen, N. Cabrol, J. Teza, P. Tompkins, C. Urmson, V. Verma, M.D. Wagner, W.L. Whittaker, "First Experiments in the Robotic Investigation of Life in the Atacama Desert of Chile," In *Proceedings of the IEEE International Conference on Robotics and Automation*, April 2005.
- [34] Eric Zbinden, Laurent Nguyen, Kurt Schwehr, "*Viz: An Automated 3D Mapping System for Planetary Exploration*," ISPRS II/I Amsterdam, Summer 2000.